NORTH-HOLLAND

The Generalized Inverse of a Sum With Radical Element: Applications

D. Huylebrouck
Aartshertogstraat 42
8400 Ostends, Belgium

Submitted by Peter Lancaster

ABSTRACT

An expression for the generalized inverse of a sum with radical element in a ring with unity is generalized to the case of a sum $\varphi + \eta$, in an additive category $\mathscr{C}$, of a morphism φ with a reflexive Von Neumann regular inverse $\varphi^{(1,2)}$ and a morphism η of $\mathscr{C}$ which is such that $1_X + \varphi^{(1,2)}\eta$ is invertible. Also, if $\mathscr{C}$ is an additive category with an involution *, φ a morphism with Moore-Penrose inverse $\varphi^\dagger$, and η such that $1_X + \varphi^\dagger\eta$, $1_X - \lambda = 1_X - (1_X + \varphi^\dagger\eta)^{-1}(1_X - \varphi^\dagger\varphi)\eta^*\varphi^{\dagger *}(1_X + \eta^*\varphi^{\dagger *})^{-1}$, and $1_Y - \mu = 1_Y - (1_Y + \varphi^{\dagger *}\eta^*)^{-1}\varphi^{\dagger *}\eta^*(1_Y - \varphi\varphi^\dagger)(1_Y + \eta\varphi^\dagger)^{-1}$ are invertible, then $\varphi + \eta - (1_Y - \varphi\varphi^\dagger)\eta(1_X + \varphi^\dagger\eta)^{-1}(1_X - \varphi^\dagger\varphi)$ has a Moore-Penrose inverse, given by $(1_X - \lambda)^{-1}(1_X + \varphi^\dagger\eta)^{-1}\varphi^\dagger(1_Y - \mu)^{-1}$. Relations with results of Wynn, Roth, and Nashed are discussed.

1. INTRODUCTION

In [12] R. Puystjens and the present author gave a necessary and sufficient condition for the existence in a ring R with identity of a Von Neumann regular inverse for $a + j$, where a is a Von Neumann regular element of R and j is an element of the Jacobson radical. The same problem was solved for the Moore-Penrose inverse in the ring R with involution. In this paper these

0024-3795/96/$15.00
SSDI 0024-3795(94)00350-M

results are generalized (Propositions 1 and 2), and it is shown that the following results follow from them:

(1) Theorems of P. Wynn on the Moore-Penrose inverse of formal power series $\sum_{\nu \geqslant 0} T_\nu z^\nu$ with $T_\nu \in M_{m \times 1}(\mathbb{C})$ for all $\nu \geqslant 0$ [or $T_\nu \in M_{1 \times n}(\mathbb{C})$ for all $\nu \geqslant 0$]. His results can be generalized from these series with vector valued coefficients to arbitrary power series $\sum_{\nu \geqslant 0} T_\nu z^\nu$ with $T_\nu \in M_{m \times n}(\mathbb{C})$ for all $\nu \geqslant 0$.

(2) A condition of Baksalary and Kala on the existence of a solution of the matrix equation $AX - YB = C$ and an equivalence condition similar to a theorem of Roth.

(3) Expressions for the least squares solutions of the equation $AX - YB = C$ (over $\mathbb{C}$).

(4) Results of Z. Nashed for perturbed matrices $A + T_p$ where $A \in M_{m \times n}(\mathbb{C})$ and $(T_p)_{p \in \mathbb{N}}$ is a sequence of matrices $T_p \in M_{m \times n}(\mathbb{C})$, for all $p \in N$, such that $\lim_{p \to +\infty} T_p = 0$.

A categorial setting (cf. [13]) is used to bring these properties together.

Definitions. Suppose $\mathscr{C}$ is an additive category, and * an involution on $\mathscr{C}$. Let $\varphi : X \to Y$ be a morphism of $\mathscr{C}$. A morphism $\chi : Y \to X$ of $\mathscr{C}$ is called an (i)-*inverse* for φ if χ satisfies the ith Penrose condition, where the four Penrose conditions are:

(1) $\varphi\chi\varphi = \varphi$,
(2) $\chi\varphi\chi = \chi$,
(3) $(\varphi\chi)^* = \varphi\chi$,
(4) $(\chi\varphi)^* = \chi\varphi$.

$\chi : Y \to X$ is called an (i, j)-*inverse* if it satisfies the ith and jth Penrose conditions, etc. The set of all (i)-inverses for φ is denoted by $\varphi(i)$, and an arbitrary element of this set by $\varphi^{(i)}$. A (1)-inverse is called a *Von Neumann* regular inverse (for short, vN regular). If a morphism $\chi : Y \to X$ exists such that all four equations are satisfied, it is unique and is called *the Moore-Penrose inverse* $\varphi^\dagger$ of φ (for short, MP inverse).

2. THE MAIN RESULTS

We note that if φ is Von Neumann regular, then $\varphi^{(1)}\varphi\varphi^{(1)}$ is a $(1, 2)$-inverse of φ. Hence, if a morphism φ admits a (1)-inverse, we may as well immediately consider a $(1, 2)$-inverse.

PROPOSITION 1. *Let $\mathscr{C}$ be an additive category. Suppose that $\varphi: X \to Y$ is a vN regular morphism of $\mathscr{C}$ (let $\varphi^{(1,2)}: Y \to X$ be any $(1,2)$-inverse of φ) and that $\eta: X \to Y$ is a morphism of $\mathscr{C}$ such that $1_X + \varphi^{(1,2)}\eta$ is invertible; let ε denote $(1_Y - \varphi\varphi^{(1,2)})\eta(1_X + \varphi^{(1,2)}\eta)^{-1}(1_X - \varphi^{(1,2)}\varphi)$. Then:*

(i) *$\varphi + \eta - \varepsilon$ is always vN regular, all its $(1,2)$-inverses being of the form*

$$\begin{aligned}&\left(1_X + \varphi^{(1,2)}\eta\right)^{-1}\varphi^{(1,2)} + \left(1_X + \varphi^{(1,2)}\eta\right)^{-1}\left(1_X - \varphi^{(1,2)}\varphi\right)\theta\\&\quad + \gamma\left(1_Y - \varphi\varphi^{(1,2)}\right)\left(1_Y + \eta\varphi^{(1,2)}\right)^{-1} - \left(1_X + \varphi^{(1,2)}\eta\right)^{-1}\left(1_X - \varphi^{(1,2)}\varphi\right)\\&\quad \times\left[\theta + \gamma - \theta(\varphi + \eta)\gamma\right]\left(1_Y - \varphi\varphi^{(1,2)}\right)\left(1_Y + \eta\varphi^{(1,2)}\right)^{-1}\end{aligned}$$

with $\theta, \gamma \in \mathscr{C}$.

(ii) *If $\tau \in (\varphi + \eta)(1)$ then $\tau \in \varepsilon(1)$.*

Proof. Note that if $1_X + \varphi^{(1,2)}\eta$ is invertible, $1_Y + \eta\varphi^{(1,2)}$ is invertible too: take $(1_Y + \eta\varphi^{(1,2)})^{-1} = 1_Y - \eta(1_X + \varphi^{(1,2)}\eta)^{-1}\varphi^{(1,2)}$. To improve legibility, let α denote $(1_X + \varphi^{(1,2)}\eta)^{-1}$, and β denote $(1_Y + \eta\varphi^{(1,2)})^{-1}$. It follows that

$$\alpha\varphi^{(1,2)}\eta = \varphi^{(1,2)}\eta\alpha;$$

$$1_X - \alpha\varphi^{(1,2)}\eta = \alpha \quad\text{and}\quad 1_Y - \beta\eta\varphi^{(1,2)} = \beta;$$

$$\varphi^{(1,2)}\beta = \alpha\varphi^{(1,2)} \quad\text{and}\quad \beta\eta = \eta\alpha;$$

$$\varphi^{(1,2)}\varepsilon = \varepsilon\varphi^{(1,2)} = 0.$$

(i): We show that $\alpha\varphi^{(1,2)} \in [\varphi + \eta - \varepsilon](1)$. The first Penrose equation is satisfied:

$$\begin{aligned}&[\varphi + \eta - \varepsilon]\alpha\varphi^{(1,2)}[\varphi + \eta - \varepsilon]\\&\quad = [\varphi + \eta - \varepsilon]\alpha\varphi^{(1,2)}[\varphi + \eta]\\&\quad = [\varphi + \eta]\alpha\varphi^{(1,2)}[\varphi + \eta] = \varphi\alpha\varphi^{(1,2)}[\varphi + \eta] + \eta\alpha\varphi^{(1,2)}[\varphi + \eta]\\&\quad = \varphi\left(1_X - \alpha\varphi^{(1,2)}\eta\right)\varphi^{(1,2)}[\varphi + \eta] + \eta\alpha\varphi^{(1,2)}[\varphi + \eta]\\&\quad = \varphi\varphi^{(1,2)}[\varphi + \eta] - \varphi\alpha\varphi^{(1,2)}\eta\varphi^{(1,2)}[\varphi + \eta] + \eta\alpha\varphi^{(1,2)}[\varphi + \eta]\\&\quad = \varphi + \eta - \eta + \varphi\varphi^{(1,2)}\eta - \varphi\varphi^{(1,2)}\eta\alpha\varphi^{(1,2)}[\varphi + \eta]\end{aligned}$$

$$+ \eta\alpha\varphi^{(1,2)}[\varphi + \eta]$$
$$= \varphi + \eta - (1_Y - \varphi\varphi^{(1,2)})\eta + (1_Y - \varphi\varphi^{(1,2)})\eta\alpha\varphi^{(1,2)}[\varphi + \eta]$$
$$= \varphi + \eta - (1_Y - \varphi\varphi^{(1,2)})\eta\{1_X - \alpha\varphi^{(1,2)}[\varphi + \eta]\}$$
$$= \varphi + \eta - (1_Y - \varphi\varphi^{(1,2)})\eta(1_X - \alpha\varphi^{(1,2)}\varphi - \alpha\varphi^{(1,2)}\eta)$$
$$= \varphi + \eta - (1_Y - \varphi\varphi^{(1,2)})\eta(\alpha - \alpha\varphi^{(1,2)}\varphi)$$
$$= \varphi + \eta - (1_Y - \varphi\varphi^{(1,2)})\eta\alpha(1_X - \varphi^{(1,2)}\varphi)$$
$$= \varphi + \eta - \varepsilon.$$

The second Penrose equation is checked similarly:

$$\alpha\varphi^{(1,2)}[\varphi + \eta - \varepsilon]\alpha\varphi^{(1,2)}$$
$$= \alpha\varphi^{(1,2)}\varphi\alpha\varphi^{(1,2)} + \alpha\varphi^{(1,2)}\eta\alpha\varphi^{(1,2)} - 0$$
$$= \alpha\varphi^{(1,2)}\varphi\varphi^{(1,2)}\beta + \alpha\varphi^{(1,2)}\eta\alpha\varphi^{(1,2)} = \alpha\varphi^{(1,2)}(\beta + \eta\alpha\varphi^{(1,2)})$$
$$= \alpha\varphi^{(1,2)}(\beta + \beta\eta\varphi^{(1,2)}) = \alpha\varphi^{(1,2)}\beta(1_Y + \eta\varphi^{(1,2)})$$
$$= \alpha\varphi^{(1,2)}.$$

Thus $\alpha\varphi^{(1,2)}$ is a particular (1, 2)-inverse of $\varphi + \eta - \varepsilon$. The general expression for $(\varphi + \eta - \varepsilon)^{(1,2)}$ is obtained by a formula given in [3].

(ii): If $\tau \in (\varphi + \eta)(1)$ then $[\varphi + \eta]\tau[\varphi + \eta] = \varphi + \eta$, or

$$[(1_Y + \eta\varphi^{(1,2)})\varphi + \eta(1_X - \varphi^{(1,2)}\varphi)]$$
$$\cdot\tau\cdot[\varphi(1_X + \varphi^{(1,2)}\eta) + (1_Y - \varphi\varphi^{(1,2)})\eta]$$
$$= \varphi(1_X + \varphi^{(1,2)}\eta) + (1_Y - \varphi\varphi^{(1,2)})\eta.$$

Multiplying on the right by $\alpha(1_X - \varphi^{(1,2)}\varphi)$ yields

$$[(1_Y + \eta\varphi^{(1,2)})\varphi + \eta(1_X - \varphi^{(1,2)}\varphi)]\cdot\tau\cdot(1_Y - \varphi\varphi^{(1,2)})\eta\alpha(1_X - \varphi^{(1,2)}\varphi)$$
$$= (1_Y - \varphi\varphi^{(1,2)})\eta\alpha(1_X - \varphi^{(1,2)}\varphi),$$

or

$$[(1_Y + \eta\varphi^{(1,2)})\varphi + \eta(1_X - \varphi^{(1,2)}\varphi)]\tau\varepsilon = \varepsilon.$$

Multiplication on the left by $(1_Y - \varphi\varphi^{(1,2)})\beta$ reduces the equation to $(1_Y - \varphi\varphi^{(1,2)})\beta\eta(1_X - \varphi^{(1,2)}\varphi)\tau\varepsilon = (1_Y - \varphi\varphi^{(1,2)})\beta\varepsilon$ or $\varepsilon\tau\varepsilon = (1_Y - \varphi\varphi^{(1,2)})\beta\varepsilon$. But $(1_Y - \varphi\varphi^{(1,2)})\beta\varepsilon = (1_Y - \varphi\varphi^{(1,2)})(1_Y - \beta\eta\varphi^{(1,2)})\varepsilon = (1_Y - \varphi\varphi^{(1,2)})\varepsilon = \varepsilon$; thus $\tau \in \varepsilon(1)$. ■

PROPOSITION 2. *Let $\mathscr{C}$ be an additive category with an involution *. Suppose that $\varphi : X \to Y$ is a morphism of $\mathscr{C}$ with MP inverse $\varphi^\dagger : Y \to X$ (with respect to the involution *) and that $\eta : X \to Y$ is a morphism of $\mathscr{C}$ such that $1_X + \varphi^\dagger\eta$ is invertible. Let ε, λ, and μ denote:*

$$\left(1_Y - \varphi\varphi^\dagger\right)\eta\left(1_X + \varphi^\dagger\eta\right)^{-1}\left(1_X - \varphi^\dagger\varphi\right),$$

$$\left(1_X + \varphi^\dagger\eta\right)^{-1}\left(1_X - \varphi^\dagger\varphi\right)\eta^*\varphi^{\dagger *}\left(1_X + \eta^*\varphi^{\dagger *}\right)^{-1}$$

and

$$\left(1_Y + \varphi^{\dagger *}\eta^*\right)^{-1}\varphi^{\dagger *}\eta^*\left(1_Y - \varphi\varphi^\dagger\right)\left(1_Y + \eta\varphi^\dagger\right)^{-1},$$

*respectively. If $1_X - \lambda$ and $1_Y - \mu$ are invertible, then $\varphi + \eta - \varepsilon$ has an MP inverse with respect to the involution *; it is the morphism $(1_X - \lambda)^{-1}(1_X + \varphi^\dagger\eta)^{-1}\varphi^\dagger(1_Y - \mu)^{-1}$.*

Proof. Write $\alpha = (1_X + \varphi^\dagger\eta)^{-1}$ and $\beta = (1_Y + \eta\varphi^\dagger)^{-1}$. We note that

$$\mu[\varphi + \eta - \varepsilon] = \beta^*\varphi^{\dagger *}\eta^*\left(1_Y - \varphi\varphi^\dagger\right)\beta[\varphi + \eta - \varepsilon]$$

$$= \beta^*\varphi^{\dagger *}\eta^*\left[\left(1_Y - \varphi\varphi^\dagger\right)\beta\eta\left(1_X - \varphi^\dagger\varphi\right) - \varepsilon\right] = 0.$$

Hence $(1_Y - \mu)^{-1}[\varphi + \eta - \varepsilon] = \varphi + \eta - \varepsilon$, and in the same way $[\varphi + \eta - \varepsilon](1_X - \lambda)^{-1} = \varphi + \eta - \varepsilon$.

In order to show that $(1_X - \lambda)^{-1}\alpha\varphi^\dagger(1_Y - \mu)^{-1}$ is the MP inverse of $\varphi + \eta - \varepsilon$, we start by checking the fourth Penrose equation:

$$\left((1_X - \lambda)^{-1}\alpha\varphi^\dagger(1_Y - \mu)^{-1}[\varphi + \eta - \varepsilon]\right)^*$$

$$= (1_X - \lambda)^{-1}\alpha\varphi^\dagger(1_Y - \mu)^{-1}[\varphi + \eta - \varepsilon],$$

or by the above property,

$$\big((1_X - \lambda)^{-1}\alpha\varphi^{\dagger}[\varphi + \eta - \varepsilon]\big)^* = (1_X - \lambda)^{-1}\alpha\varphi^{\dagger}[\varphi + \eta - \varepsilon].$$

To prove the equivalent

$$\big((1_X - \lambda)^{-1}\alpha\varphi^{\dagger}[\varphi + \eta]\big)^* = (1_X - \lambda)^{-1}\alpha\varphi^{\dagger}[\varphi + \eta],$$

we shall show that $\alpha\varphi^{\dagger}(\varphi + \eta)(1_X - \lambda^*) = (1_X - \lambda)(\varphi + \eta)^*\varphi^{\dagger *}\alpha^*$:

$$\begin{aligned}
\alpha\varphi^{\dagger}(\varphi + \eta)(1_X - \lambda^*) &= \big[1_X - \alpha\big(1_X - \varphi^{\dagger}\varphi\big)\big](1_X - \lambda^*)\\
&= 1_X - \lambda^* - \alpha\big(1_X - \varphi^{\dagger}\varphi\big) + \alpha\big(1_X - \varphi^{\dagger}\varphi\big)\lambda^*\\
&= 1_X - \alpha\varphi^{\dagger}\eta\big(1_X - \varphi^{\dagger}\varphi\big)\alpha^* - \alpha\big(1_X - \varphi^{\dagger}\varphi\big)\\
&\qquad + \alpha\big(1_X - \varphi^{\dagger}\varphi\big)\alpha\varphi^{\dagger}\eta\big(1_X - \varphi^{\dagger}\varphi\big)\alpha^*\\
&= 1_X - \alpha\varphi^{\dagger}\eta\big(1_X - \varphi^{\dagger}\varphi\big)\alpha^* - \alpha\big(1_X - \varphi^{\dagger}\varphi\big)\\
&= 1_X - \big(1_X - \varphi^{\dagger}\varphi\big)\alpha^* + \alpha\big(1_X - \varphi^{\dagger}\varphi\big)\alpha^*\\
&\qquad - \alpha\big(1_X - \varphi^{\dagger}\varphi\big),
\end{aligned}$$

and analogously for $(1_X - \lambda)(\varphi + \eta)^*\varphi^{\dagger *}\alpha^*$. The remaining Penrose conditions (1, 2, 3) are also verified by direct computation. ∎

Similar theorems can be formulated for other generalized inverses (for the group inverse, see [12]; in [7] related theorems were given for the Drazin inverse as well).

3. MATRICES OVER RINGS

Consider the category of matrices over a ring R with unity and Jacobson radical rad R. The following is a direct generalization of Propositions 1 and 2 of [12]:

PROPOSITION 3.

(i) *Suppose* $A \in M_{m\times n}(R)$ *is vN regular* (*Let* $A^{(1,2)} \in M_{n\times m}(R)$ *be any* (1, 2)*-inverse of* A), $J \in M_{m\times n}(\text{rad } R)$, *and let* ε *denote*

$$\left(1_m - AA^{(1,2)}\right)J\left(1_n + A^{(1,2)}J\right)^{-1}\left(1_n - A^{(1,2)}A\right).$$

Then

$$\varepsilon = 0_{m\times n} \quad \textit{if and only if} \quad A + J \quad \textit{is vN regular}.$$

(ii) *Suppose* $A \in M_{m\times n}(R)$ *has an MP inverse* $A^\dagger$, $J \in M_{m\times n}(\text{rad } R)$, *and let* ε' *denote* $(1_m - AA^\dagger)J(1_n + A^\dagger J)^{-1}(1_n - A^\dagger A)$. *Then*

$$\varepsilon' = 0_{m\times n} \quad \textit{if and only if} \quad A + J \quad \textit{has an MP inverse}.$$

Proof. (i): The matrices $1_n + A^{(1,2)}J$ and $1_m + JA^{(1,2)}$ are invertible, since $A^{(1,2)}J \in M_{n\times n}(\text{rad } R) = \text{rad } M_{n\times n}(R)$ and $JA^{(1,2)} \in M_{m\times m}(\text{rad } R) = \text{rad } M_{m\times m}(R)$. Now if $\varepsilon = 0_{m\times n}$, then $A + J$ is vN regular by Proposition 1(i). Conversely, if $A + J$ is vN regular, then by Proposition 1(ii), ε also has a vN regular inverse, say ζ. From $\varepsilon\zeta\varepsilon = \varepsilon$ it follows that $(1 - \varepsilon\zeta)\varepsilon = 0_{m\times n}$. But $\varepsilon\zeta \in M_{m\times m}(\text{rad } R) = \text{rad } M_{m\times m}(R)$. Thus $1 - \varepsilon\zeta$ is invertible and $\varepsilon = 0_{m\times n}$.

(ii): Again, the invertibility conditions of Proposition 3 may be dropped, because J is a matrix over the Jacobson radical. ■

As an application, we prove P. Wynn's theorems on the MP inverse of a vector with formal power series entries. Wynn (see [14]) considered the formal power series $p\{T_\nu \mid z\} = \sum_{\nu\geqslant 0} T_\nu z^\nu$, with $T_\nu \in M_{m\times n}(\mathbb{C})$ and the involution * defined by $(p\{T_\nu \mid z\})^* = \sum_{\nu\geqslant 0}(T_\nu)^+ z^\nu$, with $(T_\nu)^+$ the complex conjugate transpose of T_ν. He remarked that in general the MP inverse of such a $p\{T_\nu \mid z\}$ does not always exist, and investigated the particular cases in which $m = 1$ (or $n = 1$) and $T_0 \neq 0$.

Writing $p\{T_\nu \mid z\} = T_0 + J$ with $J = \sum_{\nu\geqslant 1} T_\nu z^\nu$, Proposition 3 provides necessary and sufficient conditions for the MP inverse of an arbitrary formal power series with $m \times n$ complex matrix coefficients to exist, and also an expression for this generalized inverse. The cases considered by Wynn are now easily derived in the particular situations of row or columm matrices.

Indeed, the necessary and sufficient condition of Proposition 3 becomes

$$\left(1 - T_0 T_0^\dagger\right)\left(1 + \sum_{\nu \geqslant 1} T_\nu z^\nu T_0^\dagger\right)^{-1} \sum_{\nu \geqslant 1} T_\nu z^\nu \left(1 - T_0^\dagger T_0\right) = 0. \qquad (¥)$$

Now if $T_0 \neq 0$, then

(1) $m = 1$ implies that $T_0^\dagger = T_0^+ (T_0 T_0^+)^{-1}$ and thus $1 - T_0 T_0^\dagger = 0$;
(2) $n = 1$ implies that $T_0^\dagger = (T_0^+ T_0)^{-1} T_0^+$ and thus $1 - T_0^\dagger T_0 = 0$.

In both cases, at least one factor in the expression (¥) is zero, ensuring in these cases the existence of the MP inverse.

4. THE MATRIX EQUATION $AX - YB = C$

In this section we discuss the existence of vN regular inverses for certain block matrices and their relation to the solution of the matrix equation $AX - YB = C$. Define $T_{m+r,\, s+n}$ as the set of matrices

$$\left\{ \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \middle| A \in M_{m \times s}(R),\, B \in M_{r \times n}(R), \quad \text{and} \quad C \in M_{m \times n}(R) \right\}$$

over a ring R. Proposition 1 can be applied by considering a sum

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} + \begin{bmatrix} 0 & C \\ 0 & 0 \end{bmatrix}$$

and supposing A and B have vN regular inverses $A^{(1,2)}$ and $B^{(1,2)}$. The $(\varphi + \eta - \varepsilon)$-morphism that was considered will now correspond to

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} + \begin{bmatrix} 0 & C \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & (1 - AA^{(1,2)})C(1 - B^{(1,2)}B) \\ 0 & 0 \end{bmatrix}.$$

Combined with some well-known methods (see [5] and [6]), it yields:

PROPOSITION 4. *Suppose A and B have vN regular inverses $A^{(1,2)}$ and $B^{(1,2)}$. Then:*

(i) *The matrix*

$$\begin{bmatrix} A & C - (1 - AA^{(1,2)})C(1 - B^{(1,2)}B) \\ 0 & B \end{bmatrix}$$

always has a vN regular inverse in $T_{s+n,\, m+r}$.

(ii) *Matrices $X \in M_{m\times m}(R)$ and $Y \in M_{n\times n}(R)$ exist such that*

$$AX - YB = C - (1 - AA^{(1,2)})C(1 - B^{(1,2)}B).$$

(iii) *Matrices $P \in T_{m+r,m+r}$ and $Q \in T_{s+n,s+n}$ exist such that*

$$\begin{bmatrix} A & C - (1 - AA^{(1,2)})C(1 - B^{(1,2)}B) \\ 0 & B \end{bmatrix} = P\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}Q.$$

COROLLARY 5. *Suppose A and B have vN regular inverses $A^{(1,2)}$ and $B^{(1,2)}$. Then the following are equivalent:*

(0) $(1 - AA^{(1,2)})C(1 - B^{(1,2)}B) = 0$.

(i) $\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$ *has a vN regular inverse in $T_{s+n,m+r}$.*

(ii) *Matrices $X \in M_{m\times m}(R)$ and $Y \in M_{n\times n}(R)$ exist such that $AX - YB = C$.*

(iii) *Invertible $P \in T_{m+r,m+r}$ and $Q \in T_{s+n,s+n}$ exist such that*

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = P\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}Q.$$

Proof. (iii) $\Rightarrow$ (i) because if

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = \begin{bmatrix} P_1 & P_2 \\ 0 & P_3 \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} Q_1 & Q_2 \\ 0 & Q_3 \end{bmatrix}$$

then

$$\begin{bmatrix} Q_1 & Q_2 \\ 0 & Q_3 \end{bmatrix}^{-1}\begin{bmatrix} A^{(1,2)} & 0 \\ 0 & B^{(1,2)} \end{bmatrix}\begin{bmatrix} P_1 & P_2 \\ 0 & P_3 \end{bmatrix}^{-1} \in \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}(1,2)$$

and it belongs to the given category of upper block triangular matrices. ■

Condition (0) was obtained by Baksalary and Kala for matrices over the complexes (see [1]).

In particular, if $B = 0$, Proposition 4 implies that the equation

$$AX = C - (1 - AA^{(1,2)})C \quad \text{or} \quad AX = AA^{(1,2)}C$$

has always a solution, and from Corollary 5 it follows that the equation $AX = C$ has a solution if and only if $(1 - AA^{(1,2)})C = 0$ or $C = AA^{(1,2)}C$. These properties are well known. They provide an additional motivation for subtracting the ε-term, as was done in Propositions 1 and 4.

As an application we focus on Roth's equivalence theorem. Roth considered matrices over a field and formulated a remarkable property of the equation $AX - YB = C$. It has been shown (see [5]) that his theorem is also true for matrices over various classes of unitary rings R:

ROTH'S EQUIVALENCE THEOREM. *If $A \in M_{m\times r}(R)$ and $B \in M_{s\times n}(R)$, R a given ring, then there exist matrices $X \in M_{r\times n}(R)$ and $Y \in M_{m\times s}(R)$ such that $AX - YB = C$ if and only if*

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

are equivalent.

R. E. Hartwig showed that the theorem does not hold over a ring R with elements a, b such that $ba = 1 \neq ab$. He considered the 1×1 matrices $A = [a]$, $B = [0]$, and $C = [1 - ab]$ over such a ring R. Then $AX - YB = C$ has no solution, but

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

are equivalent over $M_{2\times 2}(R)$.

In Corollary 5, the ring over which the matrices were considered was arbitrary. Yet, there was a condition on the matrices A and B (the vN regularity). In Hartwig's example, A and B satisfy this condition. An easy choice is $A^{(1,2)} = [b]$, $B^{(1,2)} = [0]$. Then $(1 - AA^{(1,2)})C(1 - B^{(1,2)}B) = [1 - ab] \neq 0$. Thus, $AX - YB = C$ has no solution.

Note that in this case,

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

are not equivalent over $T_{1+1,1+1}$. Thus condition (iii) of the corollary is not completely identical to the condition of Roth's theorem.

5. LEAST SQUARES SOLUTIONS

In the remaining sections, we shall only consider matrices over the complexes. If F is any complex matrix, then $\operatorname{tr} F^*F = \|F\|$ defines a norm. By a *least squares solution* of an equation $AX - YB = C$ over $\mathbb{C}$, we shall mean matrices X and Y that make the expression $\|AX - YB - C\|$ minimal (cf. [15]).

From Proposition 4 it is known that the ε-term of Proposition 1 [e.g. the expression $\varepsilon = (1_Y - \varphi\varphi^{(1,2)})\eta(1_X + \varphi^{(1,2)}\eta)^{-1}(1_X - \varphi^{(1,2)}\varphi)$] corresponds to $(1 - AA^{(1,2)})C(1 - B^{(1,2)}B)$ if the equation $AX - YB = C$ is considered. Using the norm $\|\ \|$, an interpretation for this ε-term can be given.

PROPOSITION 6. *Let* $A \in M_{m\times k}(\mathbb{C})$, $B \in M_{r\times n}(\mathbb{C})$, $C \in M_{m\times n}(\mathbb{C})$. *Then the least squares solutions of the equation*

$$AX - YB = C \tag{1}$$

are the exact solutions of the equation

$$AX - YB = C - (1 - AA^\dagger)C(1 - B^\dagger B). \tag{2}$$

Proof. In both cases, we shall prove that the general expressions for X and Y are $X = A^\dagger C - A^\dagger HB^\dagger B + (1 - A^\dagger A)W$ and $Y = -CB^\dagger + AA^\dagger CB^\dagger - AA^\dagger HB^\dagger + V(1 - BB^\dagger)$ with V and W arbitrary and H such that $AA^\dagger HB^\dagger B = AA^\dagger H$.

As in [11],

$$\begin{aligned}\|AX - YB - C\| &= \|A(X - A^\dagger C - A^\dagger YB) - (1 - AA^\dagger)(C + YB)\|\\ &= \|A(X - A^\dagger C - A^\dagger YB)\| + \|(1 - AA^\dagger)(C + YB)\|\\ &= \|A(X - A^\dagger C - A^\dagger YB)\| + \|(1 - AA^\dagger)(CB^\dagger + Y)B\\ &\quad + (1 - AA^\dagger)C(1 - B^\dagger B)\|\\ &= \|A(X - A^\dagger C - A^\dagger YB)\| + \|(1 - AA^\dagger)(CB^\dagger + Y)B\|\\ &\quad + \|(1 - AA^\dagger)C(1 - B^\dagger B)\|.\end{aligned}$$

Thus $\|AX - YB - C\|$ will be minimal iff $\|A(X - A^\dagger C - A^\dagger YB)\|$ and $\|(1 - AA^\dagger)(CB^\dagger + Y)B\|$ are zero. The general solution YB of $(1 - AA^\dagger)(CB^\dagger + Y)B = 0$ or $(1 - AA^\dagger)[YB] = -(1 - AA^\dagger)CB^\dagger B$ always exists and is given by

$$YB = -(1 - AA^{\dagger})CB^{\dagger}B + AA^{\dagger}H \tag{3}$$

for arbitrary H. This equation has a solution Y iff $[-(1 - AA^{\dagger})CB^{\dagger}B + AA^{\dagger}H]B^{\dagger}B = -(1 - AA^{\dagger})CB^{\dagger}B + AA^{\dagger}H$ or $AA^{\dagger}HB^{\dagger}B = AA^{\dagger}H$. Thus (3) has the general solution $Y = -(1 - AA^{\dagger})CB^{\dagger} + AA^{\dagger}HB^{\dagger} + V(1 - BB^{\dagger})$ for arbitrary V. The general solution of $A(X - A^{\dagger}C - A^{\dagger}YB) = 0$ or, with the value of Y that was already obtained, of $AX = AA^{\dagger}C + AA^{\dagger}HB^{\dagger}B$ always exists: $X = A^{\dagger}C + A^{\dagger}HB^{\dagger}B + (1 - A^{\dagger}A)W$. Thus the least squares solutions have the desired form.

If X and Y are solutions of (2), then

$$\|AX - YB - C\| = \|(1 - AA^{\dagger})C(1 - B^{\dagger}B)\|$$

has the minimum value and hence X and Y have the desired form. ∎

6. CONTINUITY OF VON NEUMANN REGULAR INVERSES

If a matrix $U \in M_{n\times n}(\mathbb{C})$ is invertible, and the sequence $(T_p)_{p\in\mathbb{N}}$ of matrices $T_p \in M_{n\times n}(\mathbb{C})$ is such that $\lim_{p\to+\infty} T_p = 0$, then $\lim_{p\to+\infty}(U + T_p)^{-1} = U^{-1}$. VN regular inverses do not enjoy this *continuity property* (see [4], [8], and [9]). Indeed, let $A \in M_{m\times n}(\mathbb{C})$; then $A^{(1)}$ is not necessarily unique. This yields a particular kind of problem: take for example

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad T_p = \begin{bmatrix} \dfrac{1}{p} & 0 \\ 0 & 0 \end{bmatrix}, \text{ and}$$

$$(A + T_p)^{(1)} = \begin{bmatrix} \dfrac{p}{p+1} & (-1)^j \\ 0 & 0 \end{bmatrix} \qquad (p \in \mathbb{N}_0);$$

then $\lim_{p\to+\infty} T_p = 0$ but $(A + T_p)^{(1)}$ does not converge. However, if we take

$$(A + T_p)^{(1)} = \begin{bmatrix} \dfrac{p}{p+1} & 0 \\ 0 & 0 \end{bmatrix} \qquad (p \in \mathbb{N}_0),$$

then $(A + T_p)^{(1)}$ converges to a (1)-inverse of A.

It is however possible to describe (1, 2)-inverses of a *perturbed* matrix $A + T_p$ that converge to some (1, 2)-inverse of A:

PROPOSITION 7. *Suppose* $A, T_p \in M_{m\times n}(\mathbb{C})$ $(p \in \mathbb{N})$, *and*

$$\lim_{p\to+\infty} T_p = 0.$$

There exists a sequence $(U_p)_{p\in\mathbb{N}}$ *of matrices* $U_p \in M_{m\times n}(\mathbb{C})$ *such that each* U_p *is a* (1, 2)-*inverse* $(A+T_p)^{(1,2)}$ *of* $A + T_p$, *for all* $p \in \mathbb{N}$, *and* $\lim_{p\to+\infty} U_p$ *is a* (1, 2)-*inverse of* A,

if and only if

$(1 - AA^{(1,2)})T_p(1 + A^{(1,2)}T_p)^{-1}(1 - A^{(1,2)}A) = 0$ *for all sufficiently large* $p \in \mathbb{N}$ *and some* $A^{(1,2)} \in A(1,2)$.

If, in that case, $(B_p)_{p\in\mathbb{N}}$ *and* $(C_p)_{p\in\mathbb{N}}$ *are arbitrary sequences of* $m \times n$ *matrices converging to matrices* B *and* C *respectively, the expression*

$$\begin{aligned}
&\left(1 + A^{(1,2)}T_p\right)^{-1} A^{(1,2)} + \left(1 + A^{(1,2)}T_p\right)^{-1}\\
&\quad\times(1 - A^{(1,2)}A)\,B_p + C_p\left(1 + T_p A^{(1,2)}\right)^{-1}(1 - AA^{(1,2)})\\
&\quad-\left(1 + A^{(1,2)}T_p\right)^{-1}(1 - A^{(1,2)}A)\\
&\qquad\times\left[B_p + C_p - B_p(A + T_p)C_p\right](1 - AA^{(1,2)})\left(1 + T_p A^{(1,2)}\right)^{-1}
\end{aligned}$$

provides a sequence converging to some (1, 2)-*inverse of* A.

Proof. If $\lim_{p\to+\infty}(A + T_p)^{(1,2)} = A^{(1,2)}$ for some $A^{(1,2)}$ of A, then $(A + T_p)^{(1,2)} = A^{(1,2)} + [(A + T_p)^{(1,2)} - A^{(1,2)}] = A^{(1,2)} + S_p$ with $\lim_{p\to+\infty} S_p = 0$. Since $\lim_{p\to+\infty} T_p = 0$, p can be taken sufficiently large to ensure the invertibility of $1 + A^{(1,2)}T_p$.

Write $\varepsilon_p = (1 - AA^{(1,2)})T_p(1 + A^{(1,2)}T_p)^{-1}(1 - A^{(1,2)}A)$. Thus by Proposition 1, $A^{(1,2)} + S_p \in \varepsilon_p(1)$; for $\varepsilon_p(A^{(1,2)} + S_p)\varepsilon_p = \varepsilon_p$, whence $\varepsilon_p S_p \varepsilon_p = \varepsilon_p$. Now p can again be taken sufficiently large to ensure the invertibility of $1 - \varepsilon_p S_p$. But $(1 - \varepsilon_p S_p)\varepsilon_p = 0$; thus $\varepsilon_p = 0$. ∎

EXAMPLE. Consider

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad T_p = \begin{bmatrix} \dfrac{1}{p} & 0 \\ 0 & 0 \end{bmatrix} \quad (p \in \mathbb{N}_0).$$

Using Proposition 7 for

$$A^{(1,2)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad B_p = \begin{bmatrix} b_{1p} & b_{2p} \\ b_{3p} & b_{4p} \end{bmatrix}, \quad \text{and} \quad C_p = \begin{bmatrix} c_{1p} & c_{2p} \\ c_{3p} & c_{4p} \end{bmatrix},$$

[$(B_p)_{p \in \mathbb{N}}$ and $(C_p)_{p \in \mathbb{N}}$ are arbitrary converging sequences of 2×2 matrices] yields

$$(A + T_p)^{(1,2)} = \begin{bmatrix} \dfrac{p}{p+1} & c_{2p} \\ b_{3p} & b_{3p}\dfrac{p+1}{p}c_{2p} \end{bmatrix},$$

where $(b_{3p})_{p \in \mathbb{N}}$ and $(c_{2p})_{p \in \mathbb{N}}$ are arbitrary converging sequences of scalars. This expression determines (1, 2)-inverses of $A + T_p$ converging to some (1, 2)-inverse of A.

7. CONTINUITY OF THE MP INVERSE

Let $A \in M_{m \times n}(\mathbb{C})$ be a matrix with MP inverse $A^{\dagger}$, and $(T_p)_{p \in \mathbb{N}}$ a sequence of matrices $T_p \in M_{n \times n}(\mathbb{C})$ such that $\lim_{p \to +\infty} T_p = 0$. Even for the uniquely defined MP inverse the continuity property [$\lim_{p \to +\infty} (A + T_p)^{\dagger} = A^{\dagger}$] may fail to hold. For example, if

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad T_p = \begin{bmatrix} \dfrac{1}{p} & 0 \\ 0 & 0 \end{bmatrix}, \text{ and } (A + T_p)^{\dagger} = p\begin{bmatrix} 1 & -1 \\ -1 & 1 + \dfrac{1}{p} \end{bmatrix}$$

$$(p \in \mathbb{N}_0),$$

then A and $A + T_p$ have (unique) MP inverses with respect to the conjugate-transpose involution, but $\lim_{n \to +\infty}(A + T_p)^\dagger$ does not even exist. Penrose (see [10]) has formulated a necessary and sufficient condition in relation to this problem: If $\lim_{p \to +\infty}(A + T_p) = A$ then $\lim_{p \to +\infty}(A + T_p)^\dagger = A^\dagger$ if and only if $\operatorname{rank}(A + T_p) = \operatorname{rank} A$ for sufficiently large $p \in \mathbb{N}$.

A similar statement can be formulated without using the rank of the matrices involved:

PROPOSITION 8. *Suppose* $A, T_p \in M_{m \times n}(\mathbb{C})$ $(p \in \mathbb{N})$, *and* $\lim_{p \to +\infty} T_p = 0$. *Let* $A^\dagger$ *denote the MP inverse of* A *with respect to the conjugate transpose involution* $^+$. *If*

$$(A + T_p)^\Phi = (1 - L)^{-1}\left(1 + A^\dagger T_p\right)^{-1} A^\dagger (1 - M)^{-1},$$

where

$$L = \left(1 + A^\dagger T_p\right)^{-1}(1 - A^\dagger A)T_p^+ A^{\dagger +}\left(1 + T_p^+ A^{\dagger +}\right)^{-1},$$

$$M = \left(1 + A^{\dagger +} T_p^+\right)^{-1} A^{\dagger +} T_p^+ (1 - AA^\dagger)\left(1 + T_p A^\dagger\right)^{-1},$$

then the following are equivalent:

(a) $\lim_{p \to +\infty}(A + T_p)^\dagger = A^\dagger$.

(b) $(1 - AA^\dagger)T_p(1 + A^\dagger T_p)^{-1}(1 - A^\dagger A) = 0$ *for sufficiently large* $p \in \mathbb{N}$.

(c) $(A + T_p)^\Phi = (A + T_p)^\dagger$ *for sufficiently large* $p \in \mathbb{N}$.

For any A and $T_p \in M_{m \times n}(\mathbb{C})$ with $\lim_{p \to +\infty} T_p = 0$, the above matrix $(A + T_p)^\Phi$ can be formed for sufficiently large p, and of course $\lim_{p \to +\infty}(A + T_p)^\Phi = A^\dagger$. Thus, $(A + T_p)^\Phi$ can be considered as a *modified* generalized inverse for $A + T_p$. Z. Nashed (see [8]) introduced a similar definition, and as a justification he pointed out that his modified generalized inverse coincided with the MP inverse if $A + T_p A^\dagger A = (A + T_p)A^\dagger A$ was considered instead of $A + T_p$. Here also we have that $(A + T_p A^\dagger A)^\Phi = (A + T_p A^\dagger A)^\dagger$. Indeed

$$(1 - AA^\dagger)\left(1 + T_p A^\dagger AA^\dagger\right)^{-1} T_p A^\dagger A(1 - A^\dagger A) = 0.$$

This modified generalized inverse can be used to obtain iterative formulas to compute the MP inverse of a matrix. In particular, if conditions are imposed such that $L = 0 = M$, then $(A + T_p)^{\Phi} = (1 + A^{\dagger}T_p)^{-1}A^{\dagger}$ as in the results of Nashed (see [12]) and Ben-Israel (see [2]).

EXAMPLE. If, as above, $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $T_p = \begin{bmatrix} 1/p & 0 \\ 0 & 0 \end{bmatrix}$, then

$$(A + T_p)^{\Phi} = \frac{8\left(4 + \dfrac{1}{p}\right)}{\left[\left(4 + \dfrac{1}{p}\right)^2 + \dfrac{1}{p^2}\right]^2} \begin{bmatrix} \dfrac{(2 + 1/p)^2}{2} & 2 + \dfrac{1}{p} \\ 2 + \dfrac{1}{p} & 2 \end{bmatrix} \neq (A + T_p)^{\dagger}$$

These expressions are different, and this agrees with the fact that $(1 - AA^{\dagger})T_p(1 + A^{\dagger}T_p)^{-1}(1 - A^{\dagger}A) \neq 0$. Thus $\lim_{p \to +\infty}(A + T_p)^{\dagger} \neq A^{\dagger}$, but of course, $\lim_{p \to +\infty}(A + T_p)^{\Phi} = A^{\dagger}$ is still valid.

The author would like to thank a referee for improving not only the style but also the content of the present paper.

REFERENCES

1 J. K. Baksalary and Kala, The matrix equation $AX - YB = C$, *Linear Algebra Appl.* 25 : 41–43 (1979).

2 A. Ben-Israel, A note on an iterative method for generalized inversion of matrices, *Math. Comp.* 20 : 439–440 (1966).

3 A. Ben-Israel and T. N. Greville, *Generalized Inverses, Theory and Applications*, Wiley, Chichester, 1977.

4 S. L. Campbell and C. D. Meyer, *Generalized Inverses of Linear Transformations*, Pitman, 1979.

5 R. Guralnick, Roth's theorem and decomposition of modules, *Linear Algebra Appl.* 39 : 155–165 (1981).

6 R. E. Hartwig, Roth's equivalence problem in unit regular rings, *Proc. Amer. Math. Soc.* 59(1) : 39–44 (1976).

7 R. E. Hartwig, The group-inverse of a block triangular matrix, in *Current Trends in Matrix Theory, Proceedings of the Third Auburn Matrix Theory Conference* (F. Ulhig and R. Gore, Eds.), Elsevier Science, 1987.

8 M. Z. Nashed, Perturbation and approximation for generalized inverses and linear operator equations, in *Generalized Inverses and Applications* (M. Z. Nashed, Ed.), Academic, 1976.

9 B. Noble, Methods for computing the Moore-Penrose generalized inverse and related matters, in *Generalized Inverses and Applications* (M. Z. Nashed, Ed.), Academic, 1976.

10 R. Penrose, A generalized inverse for matrices, *Proc. Cambridge Philos. Soc.* 51 : 406–413 (1955).

11 R. Penrose, On best approximate solutions of linear matrix equations, *Proc. Cambridge Philos. Soc.* 52(I) : 17–19 (1955).

12 R. Puystjens and D. Huylebrouck, Generalized inverses of a sum with radical element, *Linear Algebra Appl.* 84 : 289–300 (1986).

13 R. Puystjens and D. W. Robinson, The Moore-Penrose inverse of a morphism in additive category, *Comm. Algebra* 12(3) : 287–299 (1984).

14 P. Wynn, Upon the generalized inverse of a formal power series with vector valued coefficients, *Compositio Math.* 23(4) : 453–460 (1971).

15 K. Zietak, On a particular case of the inconsistent linear matrix equation $AX + YB = C$, *Linear Algebra Appl.* 66 : 249–258 (1985).

Received 1 July 1989; final manuscript accepted 29 November 1994